

IN THE ABSTRACT OF THE DISCLOSURE

Please amend the abstract at page 95, lines 4-20, to read as follows:

---

A13  
A computing system has a data storage device for storing a database including a classified vocabulary of terms. A processor of the apparatus is arranged to associate each term with a number of different categories of data and to associate all terms falling within the same category with a common code identifying a collocation of terms that exemplify that category so that terms in different categories are associated with different codes and can be disambiguated. The processor is arranged to write, directly or indirectly, a classified vocabulary including the terms together with the associated code onto a computer-readable storage medium or to supply an electrical signal via, for example a MODEM or a LAN/WAN. The database may be used in classification of documents, spelling checking of documents and refining of keyword search results.

---

REMARKS

Favorable reconsideration of this application as presently amended and in light of the following discussion is respectfully requested.

Claims 1, 4, 7, 8, 12, 17, 19-31, 34-36, 39, 43, 50-53, 56, 65, 70, 73-75, 79 and 80 are pending in the present application. Claims 2, 3, 5, 6, 9-11, 13-16, 18, 32, 33, 37, 38, 40-42, 44-49, 54, 55, 57-64, 66-69, 71, 72 and 76-78 have been canceled, Claims 1, 4, 7, 8, 12, 19-31, 34-36, 39, 43, 50-53, 56 and 65 have been amended and Claims 79 and 80 have been added by the present amendment.

In the outstanding Office Action, it was indicated the IDS filed October 5, 1999, failed to comply with 37 C.F.R. § 1.98(a)(1); and Claims 1, 4, 7, 12, 17, 19-31, 34-36, 39, 43,

50-53, 56, 65, 70 and 73-75 were rejected under 35 U.S.C. § 103(a) as unpatentable over Wical in view of Kishi.

Regarding the objection to the IDS, Applicants note the Information Disclosure Statement filed October 5, 1999, is the corresponding United Kingdom search report and includes the five references cited therein. It is respectfully submitted the IDS complies with all the necessary requirements and thus it is requested the references cited therein be considered and acknowledged. Alternatively, if the objection is maintained, it is respectfully requested more detailed information be provided as to why the IDS does not comply with the necessary requirements.

Claims 1, 4, 7, 12, 17, 19-31, 34-36, 39, 43, 50-53, 56, 65, 70 and 73-75 stand rejected under 35 U.S.C. § 103(a) as unpatentable over Wical in view of Kishi. This rejection is respectfully traversed.

The present invention currently includes independent Claims 1, 12, 31, 39, 65 and 75. For example, amended Claim 1 is directed to a computer processing apparatus for classifying a document including means for accessing a database structure providing a plurality of different subject matter categories. The database contains a classified vocabulary including a plurality of terms in each of the different subject matter categories with each term being classified in accordance with the subject matter category structure of the database. The computer processing apparatus also includes means for receiving in computer-readable form a text document to be classified, processor means operable to compare terms appearing in the text document with the terms in the classified vocabulary and to determine from the comparison the category for the document, and means for supplying a signal carrying data representing the text document and data associating the text document with the determined category.

On the contrary, Wical discloses a document classification system having a content processing system (see column 3, lines 29-35), which uses a knowledge base or catalog containing categories arranged in a classification hierarchy to identify or classify categories for document themes so that the content processing system maps, for each document, document themes identified in a thematic profile to categories of the classification hierarchy. The preliminary document classification profile provided by the content processing system of Wical is disambiguated using the knowledge base and a category cross-reference database (see column 4, lines 41-60). This disambiguation uses the knowledge base to analyze hierarchical relationships among the preliminarily classified categories (see column 5, lines 1-3) and utilizes the category cross-reference database to validate a classification based on the category pairs. The knowledge base includes an index table containing index information from the documents, including the text and location of the index phrases to enable identification of relationships among index heads, index context entries and themes.

Wical does not teach or suggest the claimed means for accessing a database structure providing a plurality of different subject matter categories and containing a classified vocabulary including a plurality of terms on each of the different subject matter categories with each term being classified in accordance with the subject matter category structure of the database nor the claimed processing means operable to compare terms in a text document with terms in a classified vocabulary to determine, from the comparison, a category for the text document.

Rather, Wical uses a vector modeling approach and analyzes the core content of a document using a predetermined set of categories to which further categories may be added by using a linguistic engine to expand the category set to reflect expansion of the knowledge base. Wical thus classifies documents by comparison with the categories themselves.

On the contrary, the computer processing apparatus of the present invention classifies documents not by making reference to the category structure but by comparing terms appearing in the document with terms in the classified vocabulary so that disambiguation of a term can be arrived at by the virtue of the presence within the document of a term that is included in the classified vocabulary. That is, in use of the computing apparatus defined in Claim 1, it is the terms within the classified vocabulary that the processor means uses to classify a text document. In contrast, Wical uses the categories themselves to affect classification.

Similar arguments apply to independent Claims 12, 31, 39 and 75.

Further, Claim 53 is directed to a database for use with an apparatus in accordance with Claim 1 having a database structure providing a plurality of different subject matter categories (and is also thus allowable for similar reasons as discussed above with respect to Claim 1). Further, the database contains a classified vocabulary including a plurality of terms in each of the different subject matter categories with each term being classified in accordance with the subject matter category structure of the database.

On the contrary, the knowledge catalog of Wical includes a hierarchical database of categories. The knowledge base may have an index table taken from the documents themselves, but there is no disclosure in Wical about a classified vocabulary as recited in Claim 53.

In addition, Claim 56 is directed to a database for use with an apparatus in accordance with Claim 12 having a database structure providing a plurality of different subject matter categories (and thus is allowable for similar reasons as discussed above for Claim 12). The database contains a classified vocabulary including a plurality of terms in each of the different subject matter categories with each term being classified in accordance with the

subject matter category structure of the database. The database also contains a plurality of collocations in which each collocation is associated with a specific different one of the subject matter categories and each collocation includes a plurality of terms exemplifying the associated category.

Wical does not teach or suggest a database that, in addition to providing a plurality of different subject matter categories, contains a classified vocabulary including a plurality of terms in each of the different subject matter categories, with each term being classified in accordance with the subject matter category structure let alone of a database that also contains a plurality of collocations, each associated with a specific different one of the subject matter categories and each collocation including a plurality of terms exemplifying the associated categories. Rather, the knowledge base of Wical includes subject matter categories, and may, once documents have been categorized, include an index table and an index/topic table for those documents.

In addition, independent Claim 65 includes amongst other features, storage means storing a classification scheme having a plurality of collocations, each collocation being associated with a respective different subject matter area and containing a set of terms which exemplify that subject matter area, means for comparing terms used in a document to be classified with the terms in the collocations, means for allocating the document being classified to the one of the collocations which said comparing means identifies as having the most number of terms in common with the document being classified.

On the contrary, in Wical, as discussed above, documents are classified into the knowledge base categories. Regarding the outstanding Office Action's comments concerning column 14, lines 54-66 of Wical, Applicants note this passage is concerned with the operation of the theme vector processor and refers to the fact that the co-pending Application

Serial No. 08/455,484 contains source code to generate theme strengths. The code referred to in Wical is computer software source code, that is, program instructions and is not a code representing a subject matter area. Rather, the source code is provided to execute heuristic routines to generate the theme strengths as set out in column 14 of Wical.

Kishi also do not teach or suggest the features recited in the independent claims.

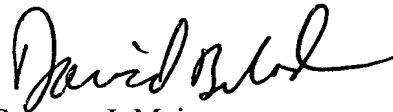
Accordingly, it is respectfully submitted independent Claims 1, 12, 31, 39, 65 and 75 and each of the claims depending therefrom are allowable.

Further, new Claims 79 and 80 have been added to set forth the invention in a varying scope, and Applicants submit the new claims are supported by the originally-filed specification. It is respectfully submitted the new claims are allowable for similar reasons as discussed above.

Consequently, in light of the above discussion and in view of the present amendment, the present application is believed to be in condition for allowance and an early and favorable action to that effect is respectfully requested.

Respectfully submitted,

OBLON, SPIVAK, McCLELLAND,  
MAIER & NEUSTADT, P.C.



Gregory J. Maier  
Attorney of Record  
Registration No. 25,599  
David A. Bilodeau  
Registration No. 42,325



**22850**

(703) 413-3000  
Fax #: (703) 413-2220  
GJM/DAB/kst  
I:\atty\DAB\105130001-am.wpd

**Marked-Up Copy**

Serial No: 09/412,754

Amendment Filed on:

March 6, 2003

IN THE CLAIMS

--1. (Amended) A computer processing apparatus for classifying a document, comprising:

means for accessing a database structure providing a plurality of different subject matter categories, the database containing a classified vocabulary [consisting of] including a plurality of terms in [all] each of the different subject matter categories with each term being classified in accordance with the subject matter category structure of the database;

means for receiving in computer-readable form a text document to be classified;

processor means operable to compare terms appearing in the text document with the terms in the classified vocabulary and to determine from the comparison the category for the document; and

means for supplying a signal carrying data representing the text document and data associating the text document with the determined category.

2. (Canceled).

3. (Canceled).

4. (Amended) [Apparatus] The apparatus according to claim 1, wherein the processor means is operable to determine the category for the document by determining from the comparison the category or categories of terms in the document, assigning weightings to the

determined categories for the terms, and assigning the document being classified to the category having the highest weighting.

5. (Canceled).

6. (Canceled).

7. (Amended) [Apparatus] The apparatus according to claim 4, wherein the processor means is operable, for each term in the classified vocabulary and in the text document, to share a predetermined weighting factor between each category associated with the term.

8. (Amended) [Apparatus] The apparatus according to claim 1, wherein the accessing means is arranged to access a plurality of collocations also forming part of the database, each collocation being associated with a specific different one on the subject matter categories and each collocation [consisting of] including a plurality of terms exemplifying the associated category.

9-11. (Canceled).

12. (Amended) A computer processing apparatus for classifying a document, comprising:

means for accessing a database having a database structure providing a plurality of different subject matter categories, the database containing a classified vocabulary [consisting of] including a plurality of terms in [all] each of the different subject matter categories with each term being classified in accordance with the subject matter category structure of the database and the database also containing a plurality of collocations each collocation being associated with a specific different one of the subject matter categories and each collocation [consisting of] including a plurality of terms exemplifying the associated category;

means for receiving in computer-readable form a text document to be classified;

processor means operable to compare terms appearing in the text document with the collocations to determine the collocation having the most terms in common with the document, and to allocate the category of the determined collocation to the document; and means for supplying a signal carrying data representing the text document and data associating the text document with the determined category.

13-16. (Canceled).

18. (Canceled).

19. (Amended) [Apparatus] The apparatus according to claim 7, wherein the accessing means is arranged to access the collocations from store means separate from the remainder of the database.

20. (Amended) [Apparatus] The apparatus according to claim 1, further comprising store means configured to store the database.

21. (Amended) [Apparatus] The apparatus according to claim 1, further comprising store means storing the database.

22. (Amended) [Apparatus] The apparatus according to claim 1, wherein the database structure provides said plurality of subject matter categories as a tree structure [consisting of] including a plurality of main subject matter areas each divided into two or more subject matter areas.

23. (Amended) [Apparatus] The apparatus according to claim 1, wherein the database structure provides said plurality of subject matter categories such that each category is defined by a subject matter area and a species or genus.

24. (Amended) [Apparatus] The apparatus according to claim 23, wherein the database provides said plurality of subject matter categories such that the species or geni are people, places, [organisations] organizations, products and technology.

25. (Amended) [Apparatus] The apparatus according to claim 23, wherein the database structure provides said plurality of subject matter categories such that the species or genus are the same for each subject matter area.

26. (Amended) [Apparatus] The apparatus according to claim 1, wherein the database provides categories in each of the following subject matter areas: the universe, the earth, the environment, natural history, humanity, recreation, society, the mind and human history.

27. (Amended) [Apparatus] The apparatus according to claim 1, wherein the database structure is such that, for a given meaning, a term is associated with only one category and different meanings of the same term are associated with different categories.

28. (Amended) [Apparatus] The apparatus according to claim 1, wherein the supplying means comprises means for storing a signal supplied by the supplying means on a computer readable medium.

29. (Amended) [Apparatus] The apparatus according to claim 1, wherein the supplying means comprises means for forwarding a signal supplied by the supplying means to another processing apparatus.

30. (Amended) [Apparatus] The apparatus according to claim 1, wherein the supplying means comprises means for displaying the information to a user.

31. (Amended) In a computer processing apparatus having means for accessing a database having a database structure providing a plurality of different subject matter categories, the database containing a classified vocabulary [consisting] including a plurality of terms in [all] each of the different subject matter categories with each term being classified in accordance with the subject matter category structure of the database and means for receiving in computer-readable form a text document to be classified, a method of classifying documents comprising:

comparing terms appearing in the text document with the terms in the classified vocabulary;

determining from the comparison the category for the document; and

supplying a signal carrying data representing the text document and data associating the text document with the determining category.

31. (Amended)

32. (Canceled).

33. (Canceled).

34. (Amended) The [A] method according to claim 31, further comprising determining the category for the document by determining from the comparison the category or categories of the terms in the document, assigning weightings to the determined categories for the terms, and assigning the document being classified to the category having the highest weighting.

35. (Amended) The [A] method according to claim 34, [which comprises] further comprising assigning weighting by, for each term in the classified vocabulary and in the text document, sharing a predetermined weighting factor between each category associated with the term.

36. (Amended) The [A] method according to claim 31, [which also comprises] further comprising accessing a plurality of collocations also forming part of the database, each collocation being associated with a specific different one of the subject matter categories and each collocation [consisting of] including a plurality of terms exemplifying the associated category.

37. (Canceled).

38. (Canceled).

39. (Amended) In a computer processing apparatus having means for accessing a database having a database structure providing a plurality of different subject matter categories, the database containing a classified vocabulary [consisting] including a plurality of terms in [all] each of the different subject matter categories with each term being classified in accordance with the subject matter category structure of the database and the database also containing a plurality of collocations each collocation being associated with a specific different one of the subject matter categories and each collocation [consisting of] including a plurality of terms exemplifying the associated category and having means for receiving in computer-readable form a text document to be classified, a method of classifying documents comprising:

comparing terms appearing the text document with the collocations to determine the collocation having the most terms in common with the document;

allocating the category of the determined collocation to the document; and

supplying a signal carrying data representing the text document and data associating the text document with the determined category.

40-42. (Canceled).

43. (Amended) The [A] method according to claim 36, [which comprises] further comprising accessing the collocations from store means separate from the remainder of the database.

44-49. (Canceled).

50. (Amended) The [A] method according to claim 31, [which comprises] further comprising carrying out the supplying by storing a signal on a computer-readable medium.

51. (Amended) The [A] method according to claim 31, [which comprises] further comprising carrying out the supplying by forwarding a signal to another processing apparatus.

52. (Amended) The [A] method according to claim 31, [which comprises] further comprising displaying the information to a user.

53. (Amended) A database for use with an apparatus in accordance with claim 1, the database having a database structure providing a plurality of different subject matter categories, the database containing a classified vocabulary [consisting of] including terms in [all] each of the different subject matter categories with each term being classified in accordance with the subject matter category structure of the database.

54. (Canceled).

55. (Canceled).

56. (Amended) A database for use with an apparatus in accordance with claim 12, the database having a database structure providing a plurality of different subject matter categories, the database containing a classified vocabulary [consisting] including a plurality of terms in [all] each of the different subject matter categories with each term being classified in accordance with the subject matter category structure of the database and the database also containing a plurality of collocations each collocation being associated with a specific different one of the subject matter categories and each collocation [consisting of] including a plurality of terms exemplifying the associate category.

57-64. (Canceled).

65. (Amended) [Apparatus] An apparatus for classifying electronic documents, comprising:

storage means storing a classification scheme having a plurality of collocations each collocation being associated with a respective different subject matter area and containing a set of terms which exemplify that subject matter area;

means for comparing terms used in a document to be classified with the terms in said collocations;

means for allocating the document being classified to the one of said collocations which said comparing means identifies as having the most number of terms in common with the document being classified;

means for associating with the document being classified a code representing the subject matter area of the allocation collocation; and

means for storing the document together with the associated code.

66-69. (Canceled).

71. (Canceled).

72. (Canceled).

76-78. (Canceled).

79-80. (New).--

#### IN THE ABSTRACT

--A computing system has a data storage device [(4, 5, 6)] for storing a database [consisting of] including a classified vocabulary of terms. A processor [(1)] of the apparatus is arranged to associate each term with a number of different categories of data and to associate all terms falling within the same category with a common code identifying a collocation of terms that exemplify that category so that terms in different categories are associated with different codes and can be disambiguated. The processor [(1)] is arranged to

write, directly or indirectly, a classified vocabulary [consisting of] including the terms together with the associated code onto a computer-readable storage medium [(RDD2)] or to supply an electrical signal via, for example a MODEM [(10)] or a LAN/WAN [(11)]. The database may be used in classification of documents, spelling checking of documents and refining of keyword search results.--